

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/532,948  
Source: Per  
Date Processed by STIC: 3/31/06

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 03/31/2006

PATENT APPLICATION: US/10/532,948

TIME: 11:37:03

Input Set : A:\2850-106.ST25.txt

Output Set: N:\CRF4\03292006\J532948.raw

3 <110> APPLICANT: Yokoyama, Shigeyuki  
 4 Shirouzu, Mikako  
 5 Sakamoto, Ayako  
 6 Sakamoto, Kensaku  
 8 <120> TITLE OF INVENTION: NON-NATURALLY-OCCURRING AMINO ACID-CONTAINING  
 PROTEIN EXPRESSION  
 9 METHOD  
 11 <130> FILE REFERENCE: P/2850-106  
 13 <140> CURRENT APPLICATION NUMBER: 10/532,948  
 C--> 14 <141> CURRENT FILING DATE: 2005-04-27  
 16 <150> PRIOR APPLICATION NUMBER: PCT/JP03/14028  
 17 <151> PRIOR FILING DATE: 2003-10-31  
 19 <150> PRIOR APPLICATION NUMBER: 2002-318846  
 20 <151> PRIOR FILING DATE: 2002-10-31  
 22 <160> NUMBER OF SEQ ID NOS: 32  
 24 <170> SOFTWARE: PatentIn version 3.3  
 26 <210> SEQ ID NO: 1  
 27 <211> LENGTH: 167  
 28 <212> TYPE: DNA  
 29 <213> ORGANISM: Artificial  
 31 <220> FEATURE:  
 32 <223> OTHER INFORMATION: an artificial base sequence consisting of a leader  
 sequence of  
 33 human tRNA gene, and the tRNA<sup>tyr</sup> gene of B. stearothermophilus  
 34 with a CUA anticodon, but without the terminal CCA sequence and a  
 35 transcription terminator  
 37 <400> SEQUENCE: 1  
 38 agcgcctccgg tttttctgtg ctgaacctca ggggacgccg acacacgtac acgtcggagg 60  
 40 ggtagcgaag tggctaaacg cggcggactc taaatccgct ccctttgggt tcggcggttc 120  
 42 gaatccgtcc cctccagac aagtgcggtt tttttctcca gtcctccg 167  
 45 <210> SEQ ID NO: 2  
 46 <211> LENGTH: 44  
 47 <212> TYPE: DNA  
 48 <213> ORGANISM: Artificial  
 50 <220> FEATURE:  
 51 <223> OTHER INFORMATION: a PCR primer used for amplifying a part of genomic  
 DNA of E. coli  
 53 <400> SEQUENCE: 2  
 54 ggaattccat atggcaagca gtaacttgat taaacaattg caag 44  
 57 <210> SEQ ID NO: 3  
 58 <211> LENGTH: 50  
 59 <212> TYPE: DNA  
 60 <213> ORGANISM: Artificial  
 62 <220> FEATURE:

63 <223> OTHER INFORMATION: a PCR primer used for amplifying a part of genomic  
DNA of E.  
64 coli.

## RAW SEQUENCE LISTING

DATE: 03/31/2006

PATENT APPLICATION: US/10/532,948

TIME: 11:37:03

Input Set : A:\2850-106.ST25.txt

Output Set: N:\CRF4\03292006\J532948.raw

```

66 <400> SEQUENCE: 3
67 gccgaagctt gtcgactttc cagcaaatca gacagtaatt ctttttaccg          50
70 <210> SEQ ID NO: 4
71 <211> LENGTH: 40
72 <212> TYPE: DNA
73 <213> ORGANISM: Artificial
75 <220> FEATURE:
76 <223> OTHER INFORMATION: a PCR primer used in the overlapping extension in
the present
77     invention
79 <400> SEQUENCE: 4
80 aggatcgaag ccgcaagcga gcgcgatcgg gccttgcgcc          40
83 <210> SEQ ID NO: 5
84 <211> LENGTH: 40
85 <212> TYPE: DNA
86 <213> ORGANISM: Artificial
88 <220> FEATURE:
89 <223> OTHER INFORMATION: a PCR primer used in the overlapping extension in
the present
90     invention
93 <220> FEATURE:
94 <221> NAME/KEY: misc_feature
95 <222> LOCATION: (16)..(16)
96 <223> OTHER INFORMATION: m represents c or a
98 <220> FEATURE:
99 <221> NAME/KEY: misc_feature
100 <222> LOCATION: (17)..(18)
101 <223> OTHER INFORMATION: n is a, c, g, or t
103 <400> SEQUENCE: 5
W--> 104 aggatcgaag ccgcammnga gcgcgatcgg gccttgcgcc          40
107 <210> SEQ ID NO: 6
108 <211> LENGTH: 33
109 <212> TYPE: DNA
110 <213> ORGANISM: Artificial
112 <220> FEATURE:
113 <223> OTHER INFORMATION: a PCR primer used in the overlapping extension in
the present
114     invention
116 <400> SEQUENCE: 6
117 acggtgtggt gctgtctatt ggtggttctg acc          33
120 <210> SEQ ID NO: 7
121 <211> LENGTH: 33
122 <212> TYPE: DNA
123 <213> ORGANISM: Artificial
125 <220> FEATURE:
126 <223> OTHER INFORMATION: a PCR primer used in the overlapping extension in
the present
127     invention
129 <400> SEQUENCE: 7
130 acggtgtggt gctggcaatt ggtggttctg acc          33
133 <210> SEQ ID NO: 8
134 <211> LENGTH: 33
135 <212> TYPE: DNA

```

## RAW SEQUENCE LISTING

DATE: 03/31/2006

PATENT APPLICATION: US/10/532,948

TIME: 11:37:03

Input Set : A:\2850-106.ST25.txt

Output Set: N:\CRF4\03292006\J532948.raw

```

136 <213> ORGANISM: Artificial
138 <220> FEATURE:
139 <223> OTHER INFORMATION: a PCR primer used in the overlapping extension in
the present
140     invention
142 <400> SEQUENCE: 8
143 acggtgtggt gctgaacatt ggtggttctg acc 33
146 <210> SEQ ID NO: 9
147 <211> LENGTH: 33
148 <212> TYPE: DNA
149 <213> ORGANISM: Artificial
151 <220> FEATURE:
152 <223> OTHER INFORMATION: a PCR primer used in the overlapping extension in
the present
153     invention
155 <400> SEQUENCE: 9
156 acggtgtggt gctgtgcatt ggtggttctg acc 33
159 <210> SEQ ID NO: 10
160 <211> LENGTH: 32
161 <212> TYPE: DNA
162 <213> ORGANISM: Artificial
164 <220> FEATURE:
165 <223> OTHER INFORMATION: a PCR primer used in the overlapping extension in
the present
166     invention
168 <400> SEQUENCE: 10
169 ttcttcggat ccaaccagac tgcgccgcct tc 32
172 <210> SEQ ID NO: 11
173 <211> LENGTH: 30
174 <212> TYPE: DNA
175 <213> ORGANISM: Artificial
177 <220> FEATURE:
178 <223> OTHER INFORMATION: a PCR primer used in the overlapping extension in
the present
179     invention
181 <400> SEQUENCE: 11
182 gatcatctgg ttaacggaga agtggtttgcc 30
185 <210> SEQ ID NO: 12
186 <211> LENGTH: 26
187 <212> TYPE: DNA
188 <213> ORGANISM: Artificial
190 <220> FEATURE:
191 <223> OTHER INFORMATION: a PCR primer used in the overlapping extension in
the present
192     invention
194 <400> SEQUENCE: 12
195 gaccttcctg tgcgatattg gcaaac 26
198 <210> SEQ ID NO: 13
199 <211> LENGTH: 12
200 <212> TYPE: DNA
201 <213> ORGANISM: Artificial
203 <220> FEATURE:
204 <223> OTHER INFORMATION: the box A consensus sequence

```

207 <220> FEATURE:

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/532,948

DATE: 03/31/2006

TIME: 11:37:03

Input Set : A:\2850-106.ST25.txt

Output Set: N:\CRF4\03292006\J532948.raw

```

208 <221> NAME/KEY: misc_feature
209 <222> LOCATION: (2)..(2)
210 <223> OTHER INFORMATION: r represents g or a
212 <220> FEATURE:
213 <221> NAME/KEY: misc_feature
214 <222> LOCATION: (5)..(6)
215 <223> OTHER INFORMATION: n is a, c, g, or t
217 <220> FEATURE:
218 <221> NAME/KEY: misc_feature
219 <222> LOCATION: (10)..(10)
220 <223> OTHER INFORMATION: n is a, c, g, or t
222 <400> SEQUENCE: 13
W--> 223 trgcnnagyn gg 12
226 <210> SEQ ID NO: 14
227 <211> LENGTH: 11
228 <212> TYPE: DNA
229 <213> ORGANISM: Artificial
231 <220> FEATURE:
232 <223> OTHER INFORMATION: the box B consensus sequence
235 <220> FEATURE:
236 <221> NAME/KEY: misc_feature
237 <222> LOCATION: (8)..(8)
238 <223> OTHER INFORMATION: n is a, c, g, or t
240 <400> SEQUENCE: 14
W--> 241 ggttcgantic c 11
244 <210> SEQ ID NO: 15
245 <211> LENGTH: 20
246 <212> TYPE: DNA
247 <213> ORGANISM: Artificial
249 <220> FEATURE:
250 <223> OTHER INFORMATION: a sequence of a primer binding site pbs1
252 <400> SEQUENCE: 15
253 agcgagtgtt aaccctgcct 20
256 <210> SEQ ID NO: 16
257 <211> LENGTH: 20
258 <212> TYPE: DNA
259 <213> ORGANISM: Artificial
261 <220> FEATURE:
262 <223> OTHER INFORMATION: a sequence of a primer binding site pbs2
264 <400> SEQUENCE: 16
265 cgactacgat attcgcgag 20
268 <210> SEQ ID NO: 17
269 <211> LENGTH: 12
270 <212> TYPE: DNA
271 <213> ORGANISM: Artificial
273 <220> FEATURE:
274 <223> OTHER INFORMATION: a sequence of a BstXI-1 site
276 <400> SEQUENCE: 17
277 ccagcagact gg 12

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/532,948

DATE: 03/31/2006

TIME: 11:37:03

Input Set : A:\2850-106.ST25.txt

Output Set: N:\CRF4\03292006\J532948.raw

```

280 <210> SEQ ID NO: 18
281 <211> LENGTH: 12
282 <212> TYPE: DNA
283 <213> ORGANISM: Artificial
285 <220> FEATURE:
286 <223> OTHER INFORMATION: a sequence of a BstXI-2 site
288 <400> SEQUENCE: 18
289 ccagcttcct gg                                     12
292 <210> SEQ ID NO: 19
293 <211> LENGTH: 63
294 <212> TYPE: DNA
295 <213> ORGANISM: Artificial
297 <220> FEATURE:
298 <223> OTHER INFORMATION: a nucleotide sequence coding a short peptide used
for
299     substitution of green fluorescent protein (cyanfluorescent
300     mutation)
302 <400> SEQUENCE: 19
303 atgggaacta gtccatagtg gtggaattct gcagatatcc agcacagtgg cggccgccgc       60
305 gtc                                                         63
308 <210> SEQ ID NO: 20
309 <211> LENGTH: 11
310 <212> TYPE: DNA
311 <213> ORGANISM: Artificial
313 <220> FEATURE:
314 <223> OTHER INFORMATION: another box B consensus sequence
317 <220> FEATURE:
318 <221> NAME/KEY: misc_feature
319 <222> LOCATION: (8)..(8)
320 <223> OTHER INFORMATION: n is a, c, g, or t
322 <400> SEQUENCE: 20
W--> 323 agttcgantc t                                     11
326 <210> SEQ ID NO: 21
327 <211> LENGTH: 31
328 <212> TYPE: DNA
329 <213> ORGANISM: Artificial
331 <220> FEATURE:
332 <223> OTHER INFORMATION: a sequence of a primer used for amplifying the
sequence of SEQ ID
333     No. 1
335 <400> SEQUENCE: 21
336 cacagaattc tcgggagctg gagaaaaaaaa c                               31
339 <210> SEQ ID NO: 22
340 <211> LENGTH: 30
341 <212> TYPE: DNA
342 <213> ORGANISM: Artificial
344 <220> FEATURE:
345 <223> OTHER INFORMATION: a sequence of another primer used for amplifying
the sequence of
346     SEQ ID No. 1
348 <400> SEQUENCE: 22
349 cacaaagctt agcgtccgg tttttctgtg                                   30

```



RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/532,948

DATE: 03/31/2006  
TIME: 11:37:04

Input Set : A:\2850-106.ST25.txt  
Output Set: N:\CRF4\03292006\J532948.raw

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220>

to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 17,18

Seq#:13; N Pos. 5,6,10

Seq#:14; N Pos. 8

Seq#:20; N Pos. 8

**Invalid <213> Response:**

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27

Seq#:28,30,31,32

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/532,948

DATE: 03/31/2006

TIME: 11:37:04

Input Set : A:\2850-106.ST25.txt

Output Set: N:\CRF4\03292006\J532948.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:104 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0

L:223 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0

L:241 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0

L:323 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0